

## IMPROVED "MIDLINK" VIRTUAL INSERTION SYSTEM AND METHODS

[0001] This application is a continuation-in-part of U.S. Patent Appln. No. 09/215,274, filed December 18, 1998, to Bruno et al., and U.S. Patent Appln. No. 09/356,382, to Bruno et al., filed July 19, 1999, the disclosures of which are herein incorporated by reference. At the time that the present invention was made, the present invention and the two applications were owned by the same entity and/or subject to an obligation of assignment to the same entity.

### FIELD OF THE INVENTION

[0002] The present invention relates to virtual insertion systems and methods for television video and, more particularly to an improvement in a "midlink" system and method which enables the virtual insertion system to be positioned downstream of the originating site in the chain of distribution video program, and wherein replacement pattern is inserted without replacing graphics previously added to the feed.

### BACKGROUND OF THE INVENTION

[0003] The term virtual insertion system is used herein to describe systems which replace, or insert in place of, in a video sequence of a scene (i.e., as obtained by a video camera or recorder), a target region or area in the video image by a matched replacement pattern adapted to be inserted into the target region, such as representation of a model stored in memory. For example, the video sequence may be of a soccer match or a stock car race wherein commercial billboards are part of the scene. The virtual insertion process involves replacement of the "target," i.e., a particular billboard advertising a first product, in the scene with a representation of a different billboard advertising a second product so that in using existing techniques, this way a different commercial

product is advertised in the scene. This can be done in such a way that the substituted portion fits substantially seamlessly into the basic image so as not to be noticeable to a viewer as a replacement or substitute.

[0004] Briefly considering existing virtual insertion systems, a representation of the target, i.e., the selected part of the scene intended for replacement or addition, is memorized, i.e., stored in memory. The position, size and perspective of the target are computed. The stored pattern is then transformed geometrically according to the estimated size and perspective of the corresponding target in the current scene image. The pattern representation is also modified in accordance with the radiographic properties of the target. Finally, the transformed pattern is inserted into the current scene image to replace the target. It will be understood that the transformed pattern need not be a sample image but can instead be a two-dimensional or three dimensional graphic element (which may or may not be animated). Systems of this general type are disclosed, for example, in U.S. Patent Nos. 5,264,933 (Rosser), 5,363,392 (Luquet et al), 5,436, 672 (Medioni et al) and 5,543,856 (Rosser) as well as French Patent No. 94-05895 to which patents reference is made for a more complete description of the virtual insertion process and the subject matter of which patents is hereby incorporated by reference.

[0005] There are two basic types of virtual insertion systems, instrumented camera systems and image recognition systems. The process used to obtain an estimation of the position, size and perspective of a target depends on whether the camera is instrumented or not. In an instrumented system, sensors are used to measure the camera operating parameters such as pan, tilt, focus and zoom, and the location, size and perspective of the target are determined from the sensor outputs. If the cameras are not instrumented and thus information from sensors is not available, an image

recognition system is used to detect and track the relevant area or areas of the current scene images in order to obtain the required parameters and the area or areas are replaced in real time.

[0006] Referring to Figure 1, the typical chain of distribution of a television program is indicated in a schematic manner. A plurality of cameras 10 are focussed on a scene S and transmit what is referred to as a "clean clean" feed, i.e., a feed without graphics or special effects, to a mobile control room or van ("truck") 12 which is generally located at the venue. Control room 12, which is generally located at the venue, i.e., at the site of the event, selects the image that will be broadcast, using a multiplexer or switcher unit. The multiplexer unit also generates a coded signal, referred to as a "tally" signal or "tally, 11 to identify the specific camera being used to produce that particular image. For economic and aesthetic reasons, only certain broadcast cameras are instrumented with sensors and the tally closure of the cameras reflects which camera is active or on air at any given time. Signals can also be generated which reflect whether a given graphic layer or special effect is in use at any given time. In the terminology generally used, a "clean-clean feed" contains only the camera signals whereas a "clean feed" contains one graphic layer and/or special effect (e.g., a slow motion replay) . Using standard video equipment, the control room can add graphic layers and/or special effects to produce the final image. A so-called "dirty feed," i.e., a feed containing the camera image plus all of the graphic layers, special effects, etc, is then sent to the network studio 16 via a satellite indicated at 14. The principal role of the network studio is to broadcast the images, via a satellite 17, to daughter stations 18 and these stations, in turn, broadcast the images to the public, as indicated by individual television receivers 19.

[0007] In some present commercial systems, cameras are used in a switched mode wherein image processing is carried out "before" the multiplexer or switcher. For example, with these prior art systems, the director in the mobile control room has two signals from camera A from which to

choose, signal A and signal A' wherein signal A' is a signal from camera A which has been previously processed at the venue and which is thus delayed with respect to signal A. There are a number of different approaches in providing virtual insertion that have been used, or are potentially useable with respect to the location at which virtual insertion takes place. A first approach, which will be referred to as an uplink monocamera system and which is illustrated in Figure 2, concerns a system or configuration wherein the virtual insertion system is located on-site, i.e., wherein the video images (and the sensor data, if applicable,) are processed locally at the venue, i.e., are sent to the mobile control room or outside broadcaster van of the broadcaster located at the venue and processed there. This is the approach typically used in some commercial virtual insertion systems.

[0008] In Figure 2, cameras 10a, 10b, 10c are connected to a multiplexer 31 and an image processing system 21 is located between the cameras and the multiplexer 31. It will be understood that Figure 2 is intended to cover the generic case, i.e., both instrumented and uninstrumented cameras, and that for instrumented cameras, both an image signal and a sensor output signal would be provided for each camera. Further, although only a single image processing system is shown, typically there would be an imaging processor for each camera. A virtual insertion device or unit 22 of the type described above replaces the relevant part, i.e., the target region, of the video image with the desired advertising pattern or the like. Again, in the commercial implementation, a separate virtual insertion unit 22 is individually associated with each camera, regardless of whether the camera is on air or not, in order to produce a different feed for use in the rest of the chain. The virtual insertion units 22 obviously must be on-site and must also be attached to each camera, where more than one camera is to be used. As mentioned above, the director in the control has the choice of two duplicate images, a "clean clean" image directly from the camera and a delayed image from the camera after processing by the image processing system 21 and the virtual insertion

unit 22, and the multiplexer 31 can be used to switch between the two. The multiplexer 31 is located in a mobile control room or van 30 along with standard video equipment indicated at 32. The video equipment 32 is used to add graphic layers, special effects and non-camera generated effects such as camera replay to the output images from the multiplexer 31. The images are sent to the network studio 40 and, from there, are relayed to daughter station (s) 50. Among the disadvantages of this approach are that one virtual insertion or replacement system is necessary for each camera and the virtual insertion operation must be performed on-site which requires that a relatively large number of technical people be on-site.

[0009] Referring to Figure 3, wherein elements corresponding to those shown in Figure 2 have been given the same reference numerals, what will be referred to as an uplink multicamera configuration is shown. In this configuration, which has been used commercially by the assignee of the present application since 1995, the virtual insertion device 22 is located in a van (e.g. an EPSIS" truck) , onsite, and accepts inputs from multiple cameras (e.g. cameras 10a, 10b and 10c) and processes the "clean feed" of the active camera, as identified by the tally signal from the mobile control room 30. In one embodiment, represented schematically in Figure 3, a pattern recognition module of the image processing system 21 is used to determine the target area to be replaced and while, in alternative embodiments, instrumented cameras are used, and data signals from camera sensors, i.e., pan, tilt, zoom and focus signals, are sent directly from the cameras 10a, 10b and 10c to the virtual insertion device 22. The modified video stream produced by virtual insertion device or system 22 is then sent back to the mobile control room or van 30 and the video equipment 32 inserts graphic layers and special effects or, alternatively, the virtual insertion device uses graphics layers and special effects from the control room 30 to generate a new "dirty feed" to the network station or studio 40.

[0010] Referring to Figure 4, a system which will be referred to as an uplink/downlink system is shown. Again, corresponding units have been given the same reference numerals as in Figure 2. In this system, the processing required, for virtual insertion is split into two parts. If image processing is to be performed, it is carried out on-site as indicated by image processing unit 21. All of the other required steps are performed at the mobile control room or van 30 except for the actual insertion. All of the information necessary to perform the insertion step (e.g., target location, occluded pixels, etc.) is encoded at the mobile control room 30 and is transmitted to the network studio 40. The virtual insertion is performed at the studio 40 or downstream thereof as indicated by the location of insertion system 22. At the daughter station(s) 50, the insertion pattern can be different for each of the daughter stations, if desired. A system of this type is disclosed in French Patent No. 94-05895, referred to above. Methods for protecting the encoded information are described in one of the above-mentioned Rosser patents (U.S. Patent No. 5,543,856) along with a "master" - "slave" system wherein the master system does the image recognition and detection and provides information pertaining to the precise location of the inserted image and the slave system carries out the insertion operation.

#### SUMMARY OF THE INVENTION

[0011] In accordance with the basic invention, a "midlink" system is provided wherein the required input and control data is collected at the venue, i.e., on-site and transported to an off-site location at which virtual insertion is performed on the "dirty feed" broadcast from the venue.

[0012] In accordance with one aspect of the basic invention a television system is provided wherein a target region in successive video images is replaced by a matching pattern adapted to be inserted into the target region, the system comprising:

at least one television camera for producing a sequence of video images of a scene;

image broadcast processing means for receiving the video images and for selectively adding layers of graphics and special effects to the video images to produce a broadcast feed; and

virtual insertion means, located off-site from the broadcast image processing means, for receiving the broadcast feed and for modifying the broadcast feed by replacing a target portion of the video images with a replacement pattern adapted to be inserted into the target portion.

**[0013]** According to a further aspect of the invention, a television system is provided wherein a target region in successive video images is replaced by a matching representation pattern adapted to be inserted into the target region, the system comprising:

at least one television camera for producing a sequence of video images of a scene;

a mobile control room located on-site with said at least one camera and including broadcast image processing means for receiving said video images and for adding layers of graphics and special effects to said video images to produce corresponding video images and means for outputting the corresponding video images in digital form as a broadcast feed; and

virtual insertion means, located off-site from said at least camera and said mobile control room, for receiving said broadcast feed and for modifying the video images thereof by replacing a target portion of said processed images with a replacement pattern adapted to be inserted into the target portion.

**[0014]** Preferably, the system includes a plurality of cameras which are adapted to be active and means for determining which one of the plurality of cameras is presently active and for producing a corresponding output, and the virtual insertion means replaces a target portion of a video image from the active camera based on said output.

[0015] Advantageously, router means are provided which are housed separately from said mobile control room and which, during calibration of the system, receive the broadcast feed and individual direct feeds from each of said cameras and selectively output one of said feeds. The router means is used to facilitate the selection of a target or targets during a calibration process for each camera prior to broadcast wherein, e.g., keyed levels are adjusted.

[0016] In the embodiment wherein the system includes a plurality of cameras, there are preferably provided means for generating camera closure signals for indicating which of said plurality of cameras is active, and means for monitoring said camera closure signals to determine if a camera closure signal has been received for the camera whose video image is currently being received by the virtual insertion means.

[0017] Preferably, the system further comprises monitoring means for monitoring the graphics and special effects added to produce the video images of the broadcast feed and for producing an output indicating that a video image received by said virtual insertion means should not be modified thereby based on the nature of the graphics and special effects that have been added to the received video image. In one preferred implementation, the monitoring means produces said output when any of the added special effects is incompatible with the replacement pattern. In a further preferred implementation, the monitoring means produces said output when any special effect has been added to the received image to be processed. Advantageously, the monitoring means produces said output when any layer of the added graphics is inconsistent with the replacement pattern.

[0018] However, often it is not necessary to prevent entirely the modification of a video image of the broadcast feed received by the virtual insertion means. Instead, the improved invention disclosed herein provides for only preventing modification of the broadcast feed in those portions of the video images where the replacement pattern would interfere with the added graphics, that is,



the replacement pattern is inserted into the video image in the target portion, except where added graphics are present. Such situations would occur, for example, when the added graphics or special effects ("graphics") are in the form of alphanumeric characters, it may be sufficient to prevent replacement of the target by the pattern only where the graphics should remain as part of the displayed image. In other words, the inserted pattern would constitute a background for the graphics.

[0019] This result can be accomplished through a number of methods. A first method comprises communicating both the clean feed and the dirty feed to the daughter stations. The daughter stations then compare the clean feed to the dirty feed to determine which parts of the video images differ. Those differing locations represent the locations of the added graphics, i.e., where the pixels of the dirty feed should not be modified. The remaining target region is replaced by the replacement pattern.

[0020] A second method is based on the use of a chroma key. In this method, the target region consists of a portion of the scene captured by the video camera that is specifically colored for that purpose (typically green or blue, in a single or multiple hues). The graphics are of a color other than the chroma key color and substitution is allowed only for the chroma key pixels.

[0021] A third method is to provide data identifying the location of the graphics in the stream of digital data sent to the daughter stations. The target region is then replaced only in those areas not occupied by the graphics, as identified. This method is suitable for use when the graphics are of a shape that can be clearly defined within the stream of digital data.

[0022] In a preferred embodiment, the at least one camera comprises a plurality of instrumented cameras each including sensor means associated therewith for producing operational data with respect to corresponding camera, and the system further comprises means for sending said

operational data to said virtual insertion means for use in replacing the target portion of a video image of the broadcast feed with a replacement pattern consistent with the operational data.

Advantageously, the operational data includes camera pan, tilt, focus and zoom.

[0023] In accordance with yet another aspect of the invention, a television system is provided wherein a target region in successive video images is replaced by a matching representation pattern adapted to be inserted into said target region, the system comprising:

a plurality of television cameras for, when active, producing a sequence of video images of a scene;

sensor means for each of said cameras for sensing a plurality of operational parameters associated with the corresponding camera and for producing a respective data output;

a mobile control room located on-site with said cameras and including image processing means for receiving said video images and for adding layers of graphics and special effects to said video images to produce resultant video images, and means for outputting the resultant video images in digital form as a broadcast feed;

local control means located on-site with said mobile control room for receiving the data outputs of said sensor means and for outputting a data signal; and

virtual insertion means, located off-site from the cameras, control room and local control means, for receiving the broadcast feed and the data signal and for modifying the video images of the broadcast feed by replacing a target portion of said video images with a replacement pattern adapted to be inserted into the target portion.

[0024] Preferably, the local control means further comprises means for determining which one of said plurality of cameras is presently active and for producing a corresponding output and the virtual insertion means receives a control signal based on that output and responsive thereto,

replaces the target portion of a video image of the broadcast feed from the camera indicated to be active.

[0025] Advantageously, the local control means further comprises router means for, during calibration of the system, receiving said broadcast feed and individual direct feeds from each of said plurality of cameras and for selectively outputting one of said feeds.

[0026] In a preferred implementation, the mobile control room includes means for generating camera closure signals for indicating which of said plurality of cameras is active and the local control means includes logic control means for monitoring said camera closure signals to determine if a camera closure signal has been received for the camera whose video image is currently being received by virtual insertion means and for sending a corresponding control signal to said virtual insertion means. Advantageously, the logic control means produces an output indicating that the second in time of two cameras is active when closure signals for a first in time camera and the second in time camera are received at the same time.

[0027] The local control means preferably further comprises logic control means for monitoring the graphics and special effects added to produce the video images of the broadcast feed and for producing an output indicating that a video image received by said virtual insertion means should not be modified thereby based on the nature of the graphics and special effects that have been added to the received processed image. As discussed above, in one embodiment, the logic control means produces said output when any of the added graphics or special effects is incompatible with the replacement pattern. Preferably, the logic means produces said output when any special effect has been added to the received video image to be processed or when any layer of the added graphics is incompatible with the replacement pattern.

[0028] Further features and advantages of the present invention will be set forth in, or apparent from, the detailed description of preferred embodiments thereof which follows.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0029] Figure 1, which was described above, is a schematic block diagram of typical chain of distribution of a television program;

[0030] Figures 2, 3 and 4, which were also described above, are block diagram representations of prior art virtual insertion systems;

[0031] Figure 5 is a block diagram representation of a virtual insertion system in accordance with the invention;

[0032] Figure 6 is a block diagram showing in somewhat more detail the on-site portion of the system of Figure 5 as implemented in accordance with a first preferred embodiment of the invention;

[0033] Figure 7 is a block diagram of a part of the on-site portion of Figure 6 as implemented in accordance with an alternative, preferred embodiment of the invention; and

[0034] Figure 8 is a block diagram of a system for providing personalized digital video compositions over a network.

#### DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0035] Referring to Figure 5, which is a schematic block diagram similar to those of Figures 2 to 4 and in which like units are given the same reference numerals, the basic elements of the midlink system of the invention are shown. In brief, as discussed above, with the system of the invention, the data is collected at the originating site and transported to a network studio, or some other

location, and processed there, off-site, in real time. In contrast to the uplink/downlink configuration of Figure 4, all of the processing necessary for the virtual insertion process is done downstream of the originating site.

[0036] In Figure 5, cameras 10a, 10b, 10c can be instrumented or uninstrumented as explained above. The video signals are sent to a mobile control room 30 which is located on-site. The control room 30 includes a multiplexer 31 used to choose which camera images are to be broadcast and conventional video processing equipment 32 used to add special effects and/or graphics to the camera images.

[0037] Once the signal processing is completed, several signals, described below, are transmitted from the mobile control room 30 to the remote location at which the virtual insertion unit 21 is located. In the exemplary embodiment illustrated, these signals are transmitted to the network studio 40 and thence to the virtual insertion unit 22 (although the virtual insertion unit 22 can, of course, be located at the network studio) . One of the transmitted signals contains information with regard to the special effects applied to the image while each of the other signals corresponds to a single layer of graphics. If the active camera is not instrumented, the parameters of the target region, i.e., the area in which the advertising is to be replaced, are obtained using a pattern recognition module of the image processing unit 21. On the other hand, if the active camera is instrumented, the data for the sensor(s) (not shown) of the camera are also transmitted to the virtual insertion unit 22. The substitution of the stored graphic pattern (e.g. advertisement) for the relevant part or target region of the video image is carried out directly by virtual insertion unit 22.

[0038] As indicated above, any one of several different configurations of the system of the invention can be employed. For example, the virtual insertion operation can be applied to the signal traversing the network control node (network studio) 40 and the resultant signal then sent to the

daughter station(s) 50. This operation can also be carried out at the daughter station(s) 50 as indicated in Figure 5. Further, a composite of these two methods can be used wherein, as indicated in Figure 5, image processing is performed on the signal received by the network studio 40 and the virtual insertion process carried out at the daughter station(s). It will be understood that multiple, different, insertions are possible at each of the daughter stations, as in the uplink/downlink system of Figure 4.

[0039] An important aspect of the present invention is that the processed images are those from a "dirty feed," i.e., a feed containing all of the graphic layers and special effects. It is noted that with multiple camera systems, the delay associated with switching between cameras, and, in particular the lack of accuracy of the tally closure delay, two tallys could close at the same time. As indicated above, the tallys or closure signals indicate which camera is on air. In the case where two tallys close at the same time, the position of the target, i.e., the location of the advertising to be inserted, could be incorrectly detected. The system of the present invention determines which camera is on--air or active using logic, i.e., a simple algorithm wherein when, e.g. . camera A has been on previously, and a tally signals are received indicating that both cameras A and B are on, it is assumed that later in time camera, camera B, is on. This simple algorithm can be implemented in hardware or software and overcomes the problems associated with the tally signal processing provided by many mobile control rooms or vans wherein the closure signal does not drop out immediately and thereby produces the ambiguity discussed above as to which camera is actually on air or active.

[0040] The present invention provides all of the advantages of the uplink/downlink system of Figure 4 but, as stated, provides offsite video stream processing downstream of the mobile control room or van. The capability of processing the video stream off-site is afforded by the provision of

the digital, as opposed to analog, transmission of the signal. This digital transmission guarantees the quality of the video stream, in contrast to an analog transmission which degrades with every satellite hop. By using the dirty feed, the images are processed downstream of multiplexer or switcher 31 and by employing the simple logic discussed above, tally synchronism problems do not occur.

[0041] In general, in accordance with the preferred embodiment, the input video signal is processed when the following condition are met: the tally of a camera is closed; no special effect is on the air, and there is no graphic layer which could affect the process (i.e., a blue graphic when the system is using a blue panel for occlusion). In this embodiment, if these conditions are not met, no processing is done.

[0042] Referring to Figure 6, further details of one preferred embodiment of the mid-link system are shown. In this embodiment, the cameras 10a, 10b and 10c are instrumented cameras as described above and have corresponding data lines 40a, 40b and 40c which are connected to a local control unit 34. The control unit 34 generally corresponds to the conventional "EPSIS'm" unit or truck which is currently used by the present assignee, (and at which the virtual insertion process is normally carried out), but performs greatly simplified functions as will become apparent. Triaxial cables 42a, 42b and 42c, which carry camera power and image signals, are connected from the respective cameras to the mobile control room or van 30. Although cables 42a, 42b and 42c can also be directly connected to local control unit 34 so as to provide a "clean clean" signal, this is not done in the preferred embodiment of the invention. A conventional multiplexer or switcher 31 is connected to conventional video equipment 32 which, as discussed above, adds whatever graphics and special effects that are to be added. It is noted that to the extent that the term multiplexer denotes or implies automatic operation, unit 31 is perhaps more accurately referred to as a switcher

or router in that, in general, the unit is selectively switched between camera outputs under the control of a director in the van.

[0043] The control room 30 also includes \* a graphic processor interface or GPI 36 which is connected to the video equipment 32 and which produces output signals based on the layers of graphics as well as special effects (e.g., slow motion) which have been added to the video image signal. These GPI signals are sent over a GPI line 46 to be input to the local control unit 34. In addition, tally signals for cameras 10a, 10b and 10c are sent to unit 34 over respective output lines 48a, 48b and 48c.

[0044] Multiplexer 31 in cooperation with video equipment 32 produces, on output line 47, a broadcast ("dirty feed") signal or feed, i.e., an edited signal which contains whatever layers of graphics and special effects that have been added to the video image signal and which is transmitted in digital form to the network studio 40. As noted elsewhere, if the control room or van 30 is adapted to produce an analog signal, conversion of the signal from analog to digital is performed. Although unit 40 is indicated to be a network studio, it would be understood that the broadcast feed can be sent to a production service studio or other type of control and then sent back to the network studio proper. Moreover, although a satellite link is illustrated in Figure 1 it will be understood that other links or pathways, such as optical links, can be used.

[0045] In the exemplary embodiment under consideration, the key component of the local control unit 34 insofar as the present invention is concerned is the router (switcher) and logic unit 38. The basic function of local control unit 34 is to make a determination as to which camera is on-air or active based on the tally signals on lines 48a, 48b and 48c and to transmit this information along with the data (sensor) signals for the active camera and the graphic and special effects information from the GPI 36. As set forth, a simple algorithm is used to determine which camera is active and a



simple logic circuit is included in router and logic unit 34 for this purpose. As was also discussed above, in accordance with a preferred embodiment, the input video signal, i.e. , the broadcast signal, is processed when certain conditions are met, viz. , the tally of the camera is closed, no special effect is on the air and there is no graphic layer which could affect the virtual insertion process. The tally signals enable the first determination to be made while the GPI signals enable the latter two determinations to be made. It is noted that decision not to process can be made at the GPI or at the local control unit 34 or even at a downstream location (e.g. at the network studio) based on the output of the switcher and logic unit 38 appearing at output line or port 49. In one embodiment, the decision to process is made at the GPI 36 which provides three different states, an "off" state wherein the GPI is turned off and two "on" states. In the "on" states, the GPI makes the determination discussed above with respect to whether the graphics or special effects are inconsistent with the processing to be done and in a first "on" state produces a "yes" or "process" signal and in the other "on" state produces a "no" or "do not process" signal. The data signals are, of course, used in determining the target area to be replaced in the virtual insertion process.

[0046] In an alternative, preferred embodiment shown in Figure 7, wherein a router or switcher 41 and a logic control unit 43 are indicated as separate elements for purposes of clarity, the broadcast signal on line 47 is sent to the router 41 of local control unit 34 and along with the camera signals, i.e., video images, on triaxial cables 42a, 42b, and 42c associated with the individual cameras 10a, 10b and 10c, respectively. As discussed above, the router 41 is used, during the calibration of the system, to switch between the individual cameras 10a, 10b and 10c and the broadcast or program feed. The output line or cable 51 is used to selectively send one of camera images or the program signal or feed to the remote location (e.g., studio 40) for processing, and a separate channel is not used for the broadcast (program) signal as in Figure 6. This remote routing of camera images

assists in calibrating the system and, in this regard, enables the chroma key to be preset. It will be understood that during calibration each camera is ranged between its extreme values for each of its operating parameters and corresponding values generated so that, for example, the camera is panned over the full range of movement thereof between its end limits, linear values are assigned the camera positions throughout the pan range and corresponding time coded signals are generated. In addition, for instrumented cameras, a preliminary calibration is performed with respect to a "snapshot" or still camera image to define or determine the target region or regions of interest. After this initial calibration step, the data signals are sent to the remote then used in the virtual insertion process as described above. The router 41 is used during program broadcasting, i.e., only the broadcast signal is sent at this time.

[0047] In the preferred embodiment of Figure 7, signals with respect to the layers of graphics and the special effects , i.e., digital video effects (e.g., fade, dissolve, slow motion, etc.) are input separately to logic control unit 43 as indicated by the separate graphics unit 53 and digital video effects (DVE) unit 55. The signal from the latter is employed as tally closure signal, i.e., a dedicated one of the tally closure pins is assigned to the DVE signal and when this pin is connected to ground, the logic unit 43 knows that the DVE is on. In contrast with the embodiment discussed above, the graphics signal is preferably a simple two state (on-off) signal.

[0048] One very important advantage of the invention is the cost savings that can be realized thereby as compared with present separate channel, commercial systems. These cost savings make the system more versatile and, in this regard, enable virtual insertion operations to be used in connection with lower rated programs (i.e. , a boxing match as opposed to the Super Bowl) These savings include those costs associated with having personnel on-site (including transportation,

lodging and other costs) that are not required with the system of the present invention because of the virtual insertion function is performed off-site.

[0049] Further, the local control unit can be highly simplified, so that the standard EPSISm mobile unit is not needed and associated costs such as those for air conditioning, power generation and a driver are saved.

[0050] It is noted that the invention is not limited, e.g. , to replacing one "billboard" with another and that the information desired to be added on be on open or barren surfaces unoccupied by elements of interest, e.g., such as on water, sand or open ground, and the like and that this information can include information about the event (e.g., for a swimming event), the names of the swimmers and the countries that the swimmers are representing can be superimposed on the swimming lanes.

[0051] Up to this point, the invention has been described when implemented in a conventional TV system. It is also suitable for use in interactive TV and for video streaming and downloading from a network such as the Internet, using IP compatible formats such as MPEG-4, Real Networks® and Windows Media Quicktime®. More generally, it can be implemented on network systems for communication with computers through methods designed to deliver media interactively or with personalization (tailoring data to the end user). Future applications of the invention include the next generation of cellular wireless telecommunication systems (UMTS, 3G wireless).

[0052] Such implementation is possible due to the high level of compression and the acceptable volume of image description data offered by standards such as MPEG-4, Digital Video Broadcasting (DVB), SML and Digital Video Multimedia Home Platform (DVD MHP). When such standards are used, packets of data describing the images may be inserted in the data stream along with the digital video data.

[0053] In a preferred embodiment of the present invention, data sufficient to describe video or graphics is placed in the data stream, thus allowing seamless integration of the additional video or graphics with the original video delivered to the end user, even in situations where there is movement of the camera and/or the presence of moving obstacles in the original video scene. A signal according to the present invention may include data describing how various still or time variable portions of an image are to be combined. Advantageously, this data may be provided by e.g., an advertiser, prior to final communication, or it may be provided in response to an action by the end user, such as by movement of the cursor to a portion of the displayed image or by clicking on a portion of the displayed image.

[0054] The present invention includes a number of methods of "personalization" of the video to the end user. First, the content owner (perhaps the company doing the video of a sporting event) may select a pre-set target region to be either retained unaltered in the display or to be substituted with a replacement object selected according to a user's profile, e.g., as defined by the language of the user's browser. Data instructions describing the location of this target region would be placed in the data stream along with the video data. In this embodiment, each receiver substitutes a selected object for the target region responsive to a comparison between stored identification data and instructions included in the feed.

[0055] In a second method, the end user (subscriber) selects a particular target region in the display to be substituted with an object. The object may be the same for all viewers, e.g., a score for a sporting event, or the object may be chosen responsive to stored information about the user, e.g., a language indication that may be determined from the user's browser.

[0056] In a third method, an object may be substituted for the target region based upon selection from a set of objects provided by a third party (e.g., an advertisement manager). Such selection

may be based on identifying characteristics such as the user's profile and the insertion may be made at any point in the feed, including at the point of identification of the target region(s).

[0057] Such processing may be made on-line (e.g., a live feed) or the resulting stream may be stored for being available on a server for services such as TV on demand. The end user can receive the program on any device able to recompose the scene from heterogeneous object sources, such as a computer, a set top box, etc.

[0058] There is no requirement that the object model to be substituted match the target zone exactly. By using chroma key ("blue screen") technology, for example, the target zone may have shape and be defined as being background while the remaining part of the image constitutes foreground. Then substitution of the pixels of a pattern larger than the target zone takes place only within in the target zone.

[0059] As an initial matter, data describing the target region must be identified. Once identified, the target description data is introduced into the compressed encoded video stream that is directed to all clients or customers. Advantageously, an image processing system, such as that disclosed in FIG. 4 (on-site) is programmed for this purpose. As noted above, the substitute object or graphics may be selected by the client, a third party, or by the end user. In addition to direct substitution, semi-transparency between the target region and the added image or graphics is achieved by further transmitting mixing factors corresponding to each image or shot during live transmission.

[0060] A system according to this aspect of the invention is shown in FIG. 8 wherein the content owner 62 provides the video stream to the server 60, typically over a network, such as the internet. While FIG. 8 shows a plurality of subservers (64, 66 and 68) all connected to server 60, and a plurality of end users (a, b and c) all connected to subserver 64, those elements shown in broken lines are for exemplary purposes only. The video stream provided by the content owner includes

therein data describing the target region(s) available for replacement as well as instructions for such replacement. In one embodiment described above, either subserver A (64) or end user a (70) selects the object to replace the target region. If the selection is made by subserver A (64), such a selection is typically based upon information known about end user a (70), such as the language of the user's browser or information supplied to subserver A (64) from end user a(70), either purposefully or automatically (e.g., through cookies).

[0061] In a second embodiment, the end user a (70) selects a particular target region in the display to be replaced. As noted above, the object may be the same for all viewers (e.g., a game score), and this object would be provided in the data stream by content owner 62, through server 60.

Another embodiment of the system comprises an advertising manager 76 that essentially supplies the replacement objects to the server to be placed into the data stream sent to the end users. Again, such objects would be chosen based either on information already known about the end users (e.g., location) or on information learned about the users. Additionally, advertising manager 76 may provide several alternative replacement objects to server 60, a selected one of which to be chosen for each target replacement by end user 70.

[0062] In an alternative embodiment, a processed video program is stored on a remote server 60 to be accessed at will via a telecommunication system using an internet-type protocol. As previously noted, the substitution of a target region can be made at the server level, according to a customer's profile. The personalized content (replacement objects) can be in the nature of still graphics, animated graphics, video, etc. Advantageously, they can be supplied in real time in the data stream, stored in a server and downloaded into a memory of a computer or set top box at times other than periods of actual use by the subscriber.

[0063] Although the invention has been described above in relation to preferred embodiments thereof, it will be understood by those skilled in the art that variations and modifications can be effected in these preferred embodiments without departing from the scope and spirit of the invention.